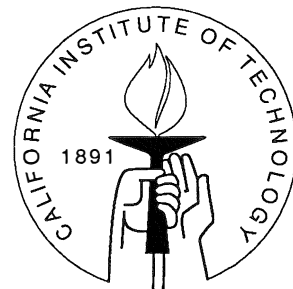


DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA 91125

Rational Voters and Strategic Voting: Evidence from the 1968, 1980, and 1992 Elections

Peter C. Ordeshook
California Institute of Technology

Langche Zeng
University of Oregon



SOCIAL SCIENCE WORKING PAPER 901

September 1994

Rational Voters and Strategic Voting: Evidence from the 1968, 1980, and 1992 Elections

Peter C. Ordeshook

Langche Zeng

Abstract

Is the rational choice paradigm more than a mere tautology when applied to the study of voting or can it generate refutable propositions that cannot be deduced or inferred from other approaches? This is the question we address empirically in the context of three-candidate presidential elections. Although we reconfirm the conclusion that the decision to vote is largely a consumptive one, we also establish that once in the voting booth, voters act strategically in precisely the ways predicted by a Downsian model of voting. That is, although expected utility calculations and the like add little to our understanding of the decision to vote, those same calculations have a significant influence on the decision for whom to vote, over and above such things as partisanship.

Rational Voters and Strategic Voting: Evidence from the 1968, 1980, and 1992 Elections*

Peter C. Ordeshook

Langche Zeng

1 Introduction

Some scholars take satisfaction from any reported increase in the number of professional papers consonant with or based on the rational choice paradigm. Others bemoan this fact. However, regardless of one's views, we cannot escape the fact that after decades of research, debate persists over the paradigm's contribution to our understanding of even the most basic elements of democratic process – voting and elections. Does the idea of rational choice offer a structure for understanding individual voting decisions and macro electoral processes in a way that allows for an empirically useful theory of elections? Or is the paradigm a mere tautology in which people vote because they are socialized to do so, in which they vote for most preferred candidates under only the most liberal interpretations of preference, and in which the derivative election models are mere mathematical toys without empirical referent?

Especially when looking at the decision to vote, we cannot deny that previous research allows only the weakest basis for arguing against a positive answer to this last question. However, although we might agree that “much of the weakness of rational choice theorizing about voter turnout ... stems from the fact that it is difficult to say what the concept of rationality brings to the explanation of the behavior in question” (Green and Shapiro, 1994, p. 98), our agreement is limited to the study of two-candidate as opposed to multi-candidate elections. Indeed, if we look more closely at what a rational choice voting calculus predicts about turnout and candidate choice in multicandidate contests, we can discern some testable hypotheses that are more than trivial or tautological.

*We would like to thank Thomas Palfrey, Morgan Kousser and Rod Kiewiet for their encouragement of this research and their suggestions, and we would especially like to acknowledge Keith Poole, not only for his comments and suggestions, but also for his assistance in organizing data from the national ICPSR election studies in an especially useful way.

To see why the predictions emanating from the paradigm are non-trivial or non-tautological when we move from two- to say, three-candidate plurality rule elections, consider the issue that lies at the heart of empirical assessments of the paradigm's usefulness – the influence (or non-influence) of an election's competitiveness. If only two candidates compete and if we model people as expected utility maximizers, then competitiveness, as a determinant of the likelihood that a person's decision is decisive for one outcome or another, can only influence a person's decision to vote. Once in the voting booth, competitiveness is irrelevant since a voter's dominant strategy is to vote for the candidate he or she most prefers. Thus, in assessing whether the rational choice paradigm brings anything useful to the study of voting in two-candidate elections when it describes people as expected utility maximizers, we are compelled to focus exclusively on the decision to vote. And it is here that we find little empirical support for supposing that the paradigm contributes much of anything.

Things are different if the number of candidates, n , exceeds two. First, the calculation of expected utility is more complicated, because the concept of an election's closeness is a multi-dimensional variable: At least in theory, if $n = 3$, then not only is the probability of a tie between a person's first and second choices (in our notation, P_{12}) relevant, but the probability of a tie between the candidates a person ranks first and third (P_{13}) and between the candidates a person ranks second and third (P_{23}) are part of that calculus as well (McKelvey and Ordeshook 1972).

Second and more interestingly, with $n = 3$, the possibility of strategic voting arises – of voting for one's second choice when a first choice is deemed uncompetitive – so that variables such as P_{12} , P_{13} , and P_{23} are not only part of the calculus describing the decision whether or not to vote, but they are also part of the calculus describing for whom to vote.¹ This fact, in turn, means that the hypothesis of expected utility maximization provides a greater variety of testable propositions, since, owing to the possibility of strategic voting, similar variables are predicted to play different roles in the decision to vote versus the decision for whom to vote. For example, as some simple algebra shows, although the probability of voting, *ceteris paribus*, ought to increase as $P_{23} B_{23}$ increases, where B_{23} is the utility differential between the candidates ranked second and third, the probability that a person votes for his or her most preferred candidate ought to decrease as this variable increases. Thus, the sign of any econometrically estimated parameter associated with $P_{23} B_{23}$ should depend on whether our dependent variable concerns the decision to vote or for whom to vote.

In addition, different measures of closeness can play different roles, depending on our dependent variable and the population subsample from which our data is drawn. For

¹See Duverger, 1951; Downs, 1957; Farquharson, 1969; Myerson and Weber, 1988; and Palfrey, 1989. For empirical assessments of such voting that are precursors to this study, see Black, 1978; Cain, 1978; Laver, 1987; Kiewiet 1979; Abramson, Aldrich, Rhode, 1982; and Abramson, Aldrich, Paolino, and Rhode, 1992

example, as we show later, when looking at candidate choice among voters who prefer a minor party candidate, the probability that a person votes for a most preferred candidate ought to decrease as $P_{23}B_{23}$ increases, whereas the influence of $P_{12}B_{12}$ and $P_{13}B_{13}$ ought to be insignificant (though otherwise positive). In contrast, if the identity of the least viable candidate is uncontroversial, voters who most prefer a viable candidate should not vote strategically for a second choice, and, therefore, should not be influenced by the PB terms.

We hasten to add that in assessing the paradigms' potential contribution, the role we predict for different variables depends on how we formulate a person's calculus. Focusing for a moment on the decision to vote, we can identify three distinct models, where the first two are theoretical competitors and the third derives from the way the Downsian model is operationalized for testing. The first, implicit thus far in our discussion, is the expected utility maximization or Downsian model (Downs, 1957; Riker and Ordeshook, 1968), which states that citizens vote or abstain depending on the expected utility of voting versus abstaining. Aside from variously postulated private benefits and costs, the key variable here is PB , where P is the probability that the citizen can influence the election outcome, and B is the utility gain (loss) associated with being decisive. The second model is the "minmax regret" one (Ferejohn and Fiorina, 1974, 1975), which posits that people choose their actions so as to minimize their maximum regret.²

These two approaches have different testable consequences. With respect to turnout, the Downsian model implies that, if utility differences between the candidates influence decisions at all, they do so through their interaction with the various measures of an election's competitiveness. The regret model implies that only utility differences matter. And with respect to the candidate choice decision, moving from two- to three-candidate elections affords us additional opportunities to discriminate between these two approaches. The Downsian model allows strategic voting — voting for one's second choice when a first choice is unlikely to win — while the minmax regret model predicts that voters always choose their first-ranked alternative (for part of the debate over these two models see Rosenthal and Sen 1973; Ferejohn and Fiorina 1975; Ashenfelter and Kelly 1975; Black 1978; Foster 1984; Hansen, Palfrey and Rosenthal 1987). Empirical applications of the Downsian model implicitly introduce a third. Because that model is often taken to mean simply that "closeness matters," it is not uncommon for research to focus only on the P term rather than on the interaction of closeness and utility (see, for instance, Ashenfelter and Kelly 1975; Black 1978). However, a model in which P but not PB influences turnout is distinct from an *expected* utility maximization model, and any inferred empirical relevance of "closeness" cannot be taken as support for the Downsian model. Such findings might support the hypothesis that candidates spend greater resources mobilizing

²"Regret" is a person's utility loss resulting from an action that may not be optimal given the "true" but unknown state of the world. The loss may be different under different "true states," and a person is assumed to choose an action such that the maximum regret associated with this action is the minimum among all acts.

citizens to vote in close elections, in which case such studies should be treated less as a test of the Downsian approach and more as an indication of the need to reconceptualize what it is we think people maximize in elections and what motivates people to vote.

In addition to trying to test the non-trivial and non-tautological hypotheses of rational choice in 3-candidate contexts, this essay also addresses some methodological issues. Specifically, because the variables specified by each of these models are, by construction, closely correlated — the PB term of the Downsian model is the product of P of the closeness model and B of the minmax regret model — all of the problems associated with multicollinearity are likely to rear their head and confound our analysis (Foster 1984; Ferejohn and Fiorina 1975). Adding to our problems is the fact that, rather than employing a standard linear model, we want to use a logit model since our dependent variables are discrete. Unfortunately, standard treatments of multicollinearity are less well-developed in this instance. Thus, we later offer a brief digression about the treatment of this problem in our analysis.

Methodological issues aside, the central question we address here is whether a simple rational-actor model can make verifiable predictions about voting that go beyond the trivial or nearly tautological propositions commonly associated with it such as: Vote for the candidate you most prefer; or vote if the cost of voting does not exceed benefits. Here, then, in order to assess whether the rational choice calculus has any empirical power, we offer a reconsideration of that calculus in the context of the three most recent U.S. Presidential elections with a “meaningful” third-party candidate – 1968, 1980, and 1992. In general, our findings bolster the conclusions Abramson *et al* (1992) offer about voting in presidential primaries: voters are strategic or at least a sufficient number of them are strategic so as to make an expected utility model of voting a relevant part of our explanation for voting patterns in multi-candidate contests. In the next section, Section 2, we discuss the specification of rational turnout models, and present empirical findings on the comparative merits of these models; Section 3 focuses on the voter’s choice and strategic voting; and Section 4 offers concluding remarks. Technical details about hypothesis testing are contained in the appendix.

2 The Decision to Vote

Downs’s (1957) expected utility model views voting as an investment decision in which a person compares the contemporaneous costs of voting with the *expected* utility gain in the future (which may be based on retrospective evaluations) and votes if the expected benefit exceeds costs. In 2-candidate elections, the net return of voting is expressed as:

$$R = PB - C \tag{1}$$

where P is the subjective probability that the voter can make a (favorable) difference in the election outcome and which is assumed to be positively related to the objective

closeness of the race. The variable B is the benefit the voter receives from this difference, and C is the cost of voting.

The difficulty here is that if $C > 0$, expression (1) fails to provide a satisfactory model of turnout in mass elections, because the probability that a single vote can make a difference is practically zero and hence the PB term is infinitesimal. So, barring the possibility that voters seriously overestimate P , it is irrational for people with any significantly positive voting cost to vote. In response, Riker and Ordeshook (1968) introduce a D term into the calculus, which is intended to represent the positive satisfaction voters associate with compliance with the ethic of voting (“citizen duty”). Because the magnitude of D is presumed to be substantial, its inclusion renders voting rational if only as a consumptive rather than investment act. Equation (1) then becomes:

$$R = PB + D - C \quad (2)$$

Insofar as multicandidate contests are concerned, if only two-way ties are likely, expression (2) generalizes to (McKelvey and Ordeshook 1972):

$$E^k - E^0 = B_{k1}P_{k1} + B_{k2}P_{k2} + \dots + D - C \quad (3)$$

$E^k - E^0$ is the expected utility difference between voting for candidate k and abstaining, and the BP terms on the right hand side represent pairwise comparisons between candidate k and the other candidates.

It is, of course, an understatement to say that the inclusion of D as a parameter of a citizen’s voting calculus is controversial. It appears that voting has been rendered rational merely by assumption, thereby revealing the rational choice paradigm as a potentially vacuous tautology able to fit itself to any data and, thereby, not subject to falsification. Unsurprisingly, numerous efforts have been made to render the act of voting theoretically rational in “less ad hoc ways” (see, for example, Palfrey and Rosenthal 1985, Ledyard 1984, Schwartz 1987). These efforts have been largely unsuccessful, but the simplest and seemingly the most fruitful approach is the one offered by Ferejohn and Fiorina (1974, 1975). Assuming that the probability terms are unknown or unknowable, Ferejohn and Fiorina consider an alternative decision rule, the minmax regret rule, which specifies that a person should choose the act that minimizes maximum regret. In this case only the B terms enter the calculation of the turnout decision, and voting is presumed to be rational if a person perceives a sufficiently great difference among the candidates, regardless of the election’s competitiveness.

Closeness Matters. So? Although the empirical significance of the D term seems well established (Riker and Ordeshook 1968; Katosh and Traugott 1982; and Ashenfelter and Kelly 1975), as long as the importance of the PB term remains in doubt, the relevance of the rational choice paradigm’s sometimes obtuse mathematical manipulations and

recourse to notions of expected utility remain in question. Indeed, we are unaware of any strong, unambiguous empirical support for the significance of PB in mass elections. There are, of course, a number of studies that find that “closeness matters”, *i.e.*, closeness alone as an independent variable matters (see, for example, Cohen and Uhlaner 1991; Foster 1984; Palfrey and Rosenthal 1987; Rosenthal and Sen 1973). But does the fact that “closeness matters” establish empirical support for the Downsian model? In other words, is “closeness matters” equivalent to “expected utility terms matter”? If not, what do these empirical findings suggest?

Recall that “expected utility” refers to the utility derived from possible future benefits (and/or costs), or more accurately, utility of some *lotteries* (*c.f.*, Varian 1984). Roughly speaking, to maximize expected utility means to choose the action that generates the greatest future net benefits *discounted* by the probabilities that these benefits will be realized. Note that the main body of the object of maximization is benefits and costs, and that probabilities of the states of the world only serve as discount factors, which generate no utility on their own. Voters get no direct utility from breaking ties, but only from the benefits they receive when their candidate wins and offers programs they favor. Hence, to show that voters are maximizing *expected utility*, we need to show that the discounted future benefits (*i.e.*, PB) matter, not that the discount factor (P) itself matters. The variable P does not constitute an expected utility term, and if it is significant in its own sake, it must be that it bears its own substantive meaning and is not playing its role as a probability discount factor. That is, if P but not PB enters the calculus significantly, then the voter cannot be said to be maximizing expected utility.

Since research affirming that “closeness” but not PB matters does not actually support the Downsian model to the exclusion of other possibilities, we offer here an formulation of an alternative model that is distinct from the Downsian and Ferejohn and Fiorina formulations. Specifically, we can let:

$$R = f(P) + D - C \quad (4)$$

or, for multi-candidate competitions,

$$R = f(P_{ij}, \forall i \neq j) + D - C \quad (5)$$

where f is some (possibly increasing) function (that we assume is linear in our empirical analysis).³

³Theoretically at least, the Downsian model and the P model intersect, since closeness may generate utility in and of itself to the extent that it measures the effort of candidates to “get out the vote.” Hence, we can include the P terms in the Downsian framework and write:

$$E^k - E^0 = B_{k1}P_{k1} + B_{k2}P_{k2} + \dots + f(P_{ij}, \forall i \neq j) + D - C \quad (6)$$

where $f(P_{ij}, \forall i \neq j)$ is but an ancillary measure of D . Later we consider this possibility in more detail, but for the present we want to keep our analysis of P , B , PB , and D separate.

As for the substantive meaning of P in (4) or (5), Rosenthal and Sen (1973) hypothesize that the level of the competition is positively correlated with a voter's interest in a campaign, which in turn stimulates turnout. We conjecture that closeness of the race is also positively related to a voter's feeling of efficacy when voting. Note that compared with its role in expected utility considerations, the closeness of the race need not to be very high to stimulate interest or to introduce efficacy. In Ferejohn and Fiorina's language, "for expected utility maximization the probability...is very important," whereas as a stimulating source, "the mere logical possibility of such an event (that an individual has any impact) is enough" (Ferejohn and Fiorina, 1975).

We have, then, three alternative models, and one way to evaluate them empirically is to include the variables specified by each model into one general equation in order to see which are significant and which are not (see, for example, Ashenfelter and Kelly 1975; Ferejohn and Fiorina 1975). In doing so, however, we encounter a serious problem with multicollinearity.

Multicollinearity Multicollinearity arises when two or more independent variables or combinations of variables are highly but not perfectly correlated. For the linear regression model, the effects, diagnosis, and remedy of this problem are well known (see, for example, Judge *et al* 1985). Although little work has been done for non-linear models in general, recent advances do offer information about the most widely used model in voting studies, binary logit (*e.g.*, Marx and Smith 1990; Schaefer 1986; Schaefer, Roi and Wolfe 1984). The effect of multicollinearity here is found to be similar to that in the linear regression: it causes extreme sensitivity of parameters to small perturbations in the explanatory variables, large variances of coefficients and predictions, and bad properties of statistical testing on parameters. Thus, diagnosis of the problem where it is likely to be present is necessary, and actions should be taken accordingly.

Some useful diagnostic tests are offered by Marx and Smith (1990), and here we briefly discuss the ideas behind them. Note that the covariance matrix of the coefficients estimated by the maximum likelihood procedure is approximated by the inverse of the (estimated) information matrix, hence the properties of the information matrix will directly affect that of the estimators. If it is ill-conditioned, *i.e.*, if it is near singular, or if its square root has nearly linearly dependent columns⁴, then the just noted undesirable properties are present. Denote the estimated information matrix by \hat{F} , and its square root by S . Denote S^* the scaled S such that its columns have unit length, and denote $F^* = S^{*'}S^*$. Then, diagnosis of ill-conditioning of \hat{F} can be based on the following:

1. The correlations between columns of S : high correlations indicate ill-conditioning of \hat{F} . The correlation matrix is given by F^* .

⁴The square root of a matrix A is a matrix D such that $D'D = A$.

2. The variance inflation measures: diagonal elements of the inverse of F^* give a measure of inflation of the asymptotic variances of the estimated parameters. These measures are 1 when the columns of S are perfectly orthogonal (ideal condition).
3. The condition numbers of F^* : defined as the square root of the ratio of the largest eigenvalue of F^* over the others, indicate ill-conditioning of F if any of them is “large”.⁵
4. Variance proportion: each eigenvalue of F^* contributes to the variance of the estimated coefficients. If the proportion of the variance of a coefficient attributed to a small eigenvalue (associated with a large condition number) is big (say, exceeds 50 percent), that coefficient is likely affected by multicollinearity, and its value is unreliable.

To show that multicollinearity exists when alternative models are compared by including all variables in the same equation, we apply the above tests to a model that includes PB , B , and P terms, as well as some variables assumed to be indicators of D and C . Specifically, the model is:

$$\begin{aligned} \text{logit}(p[\text{vote}]) = & \{Const., PB_{12}, PB_{13}, PB_{23}, P_{12}, P_{13}, P_{23}, B_{12}, B_{13}, \\ & U_1, Info., Educ., Party\} * \beta. \end{aligned} \quad (7)$$

where⁶

logit specifies the logit operation (since our dependent variable “voted-didn’t vote” is discrete),

$p[\text{vote}]$ is the probability of voting,

Const. is a constant term,

P_{ij} denotes the closeness of the race between the respondents’ i th and the j th ranked preference. Absent a cardinal subjective probability index, the actual vote return in the respondent’s state is used to construct the closeness terms.⁷ Following Myerson

⁵There is no strict criteria as how large is “large”. Johnston (1984) sets the critical value to 20, Marx and Smith to 30. Our own experience suggests that caution should be applied when any conditional number exceeds, say, 10.

⁶Note that the multi-candidate models are used here because all our data is taken from elections with three presidential candidates. Since the three B terms are perfectly collinear ($B_{13} = B_{12} + B_{23}$), only two of them are included. It does not matter which two to include, because the third provides no information.

⁷The use of the state level data is based on two considerations. The first is the simple pragmatic fact that we need a variable that in fact varies across respondents. The second reason is that it is possible to argue the special relevance of the state-wide contest given the electoral college (see, *e.g.*, Kiewiet 1979).

and Weber (1988), we operationalize the probability of a tie between the j^{th} and the k^{th} preference as: $P_{jk} = v_j * v_k$, where v_j denotes the vote percentage received by the voter's j^{th} preference which is positively related to the marginal probability that his j^{th} preference is in first place.

B_{ij} denotes the utility difference between the respondent's i th and j th ranked preference. Preference orderings are obtained from the thermometer scores, with ties broken by party scores.

$$PB_{ij} = P_{ij} * B_{ij},$$

U_1 is the utility for the respondent's most preferred candidate,⁸

Info. is an index of voter information (see Palfrey and Poole 1987 for details on the construction of this index),

Educ. education level, and

Party is strength of party identification.

The last four variables are proxies of the D and C terms.

A first sign of multicollinearity when estimating model (7) is a dramatic change in the estimated parameters when some terms are dropped from the full model. For example, in estimating the 1968 sample, the full model estimates show that the PB terms are significant. However, if the P and B terms are dropped, then PB is no longer so. Formal test results (not included here) shows that severe multicollinearity exists for all years, with the estimates of the coefficients for P and for some of the PB and B terms most likely to be affected. The estimates of the coefficients of the P terms are especially unreliable, given that more than 90% of their variances are attributed to the smallest eigenvalue (associated with the least important eigenvector), and these variances are highly inflated. The maximum correlation coefficients are well above 0.5, and the conditional numbers are close to 20. To see whether the problem exists for a simple model such as the one specified by Ferejohn and Fiorina (1975), we can estimate a model in which only the constant, one PB and one B term are included (here we include PB_{12} and B_{12}). The results are no less striking, especially for the year 1980. The condition number of the information matrix is as high as 83; the variances of the estimated coefficients are extremely inflated and are almost solely attributed to the smallest eigenvalue. Hence, the estimation of this model does not provide *any* information.

⁸ U_1 is included to test the alienation hypothesis (Riker and Ordeshook 1968). The prediction is that the higher U_1 , the more likely individuals will vote, as they feel more positively about the alternatives they face.

It is evident, then that we can not base a comparison of our three models on the MLE of model (7). We have two alternatives. First, we can derive alternative estimators for model (7). Second, we can estimate the various models separately, and perform non-nested hypotheses testing. Elsewhere we describe the results of the first approach (Zeng 1993) and here, for brevity, we focus on the second.

Empirical results. The four models we estimate separately are the following:

1. the Downsian model, obtained by dropping the P and B terms from (7);
2. the closeness model (the P-model), obtained by dropping the PB and B terms from (7);
3. the minmax regret model, obtained by dropping the PB and P terms from (7); and,
4. a virtual “social-psychological” model formed by dropping all the PB , P and B terms from (7).

The results are presented in Tables 1-1 (1968), 1-2 (1980), and 1-3 (1992). The Downsian model, the P-model, and the minmax regret model are not nested (neither is a parametric special case of the other), so classical tests such as the likelihood ratio test are not appropriate (the likelihood ratio test is applied in comparing the fourth model with others). Instead, non-nested hypothesis testing should be performed, and in the appendix we offer a convenient critical value table for one such test. To illustrate the use of this table, take the comparison between the P- model and the Downsian model for 1968. The two models have the same number of coefficients, but the absolute value of the log likelihood of the P-model is 4.06 lower than of the Downsian model ($l_1 - l_2 = 4.06$), which exceeds the critical value of 3.32 for the P-model to be judged better with probability 99%. Similarly, the P- model outperforms all others with a probability of at least 90%, whereas the Downsian model is a “Condorcet loser.” The minmax regret model is slightly better than the social-psychological model (model 4) only for 1968. With respect to the other variables in the analysis, *Info.*, *Educ.* and *Party* are highly significant, which is consistent with the findings of previous research that social-psychological factors are important predictors of the decision to vote. U_1 (*alienation*) and B_{13} (*indifference*) matters in 1968, but 1980. The PB terms are not significant for any year, whereas the P terms are significant – thereby indicating that the Downsian model and the P-model are substantially, as well as substantively different.⁹

⁹Note that our estimates for individual coefficients of the P terms are not reliable because tests show that they are seriously affected by multicollinearity. However, the precision of these estimates is an unimportant issue here since we are mainly interested in whether the P terms matter as a whole group,

Thus, it is the closeness of the race itself that affects turnout rather than some expected utility calculation. We cannot, then, reject the argument that the Downsian model fails as a useful model of the decision to vote and that any influence of "closeness" is attributable to some factor not explicitly in the model that stimulates citizens to go to the polls when elections are contestable – the P terms are actually part of some more general D term.¹⁰

3 Vote for Whom? — Strategic Voting

Were we to stop here, the criticisms offered by Green and Shapiro (1994) and others about the deficiencies of research within the rational choice paradigm would once again be reaffirmed, for we would not have added much to the literature other than support for the idea that "closeness" influences D through, perhaps, the efforts of candidates to mobilize their support when mobilization matters. On the other hand, suppose we find, using the same data, that the variables which fail to account for variations in turnout account for the decision for whom to vote? In this event, the addition of D terms and the like to a citizens voting calculus can be viewed as less an ad hoc attempt to salvage the paradigm from an embarrassing empirical fact (non-zero turnout) and more as a straightforward finding that voting is largely a consumptive rather than a strategic act, but that once in the voting booth, those same people act strategically and in accord with non-tautological notions of preference and utility. What has made such arguments difficult if not impossible to sustain in previous research is the extent to which that research focuses on two-candidate contests, or on the actions of voters with respect to the two strongest candidates. As we note earlier, P terms or PB terms are irrelevant to a voter's decision if only two candidates compete since that decision should be: Vote for the most preferred candidate. But things are different in three-candidate contest, since here a voter must be concerned that to vote for one's most preferred candidate is equivalent to throwing one's vote away or, worse still, equivalent to voting for one's least preferred candidate.

Insofar as our three models is concerned, the minmax regret model predicts that voters

and statistical tests on the significance of the group is unaffected by the precision of each parameter within it. The comparison among different models is valid, then, and we must conclude that in predicting turnout at least, the P model outperforms the other models, although precise estimates of the parameters are unknown.

¹⁰If we conclude that P merely adds to whatever else influences a citizen's "sense of duty," then we can justify including the P terms along with our indicators of D , *alienation*, *information*, *education* and *party* (see footnote 3). This, in turn, suggests estimating models that include P , D and PB as well as P , D and B . Aside from contending with the matter of multicollinearity, the resulting estimates leaves our conclusions largely unchanged. For all three years it is the P and D terms, but not PB that are significant, and a model with P and D terms alone remains superior to all others. The sole exception to this generalization is 1992, in which the coefficient for B_{12} is significant and a model including P , D , and B_{12} terms is slightly better than the rest.

never vote for their second choice (Ferejohn and Fiorina 1974). The P-model provides no prediction whatsoever: it states only that closeness affects the turnout decision in much the same way as the D term, which generates the same consumption value of *voting itself* regardless of how a person votes. The Downsian model, however, does provide some theory for the candidate choice decision. Specifically, consider the vote choice between first and second preferences.¹¹ From equation (3) or (6), the utility difference between voting for a first versus a second preference is (for the algebra, see Black 1978):

$$E^1 - E^2 = 2P_{12}B_{12} + P_{13}B_{13} - P_{23}B_{23} \quad (8)$$

and an expected utility maximizing voter will vote for his first preference rather than his second only if $E^1 - E^2 > 0$. For example, then, if the first-ranked candidate is not viable— if P_{12} and P_{13} are small— then the voter is more likely to vote for the second choice, especially when the race between the second and third ranked candidates is close and the utility difference between them is substantial (so that $P_{23}B_{23}$ is high).

Testing the Downsian model when there are more than two candidates and, thereby, testing the notion of “strategic voting”, is itself important insofar as the notion of strategic voting provides a basis for theorizing about the number of parties and the influence of winner take all election rules on that number (Duverger 1967, Palfrey 1989). A cursory review of the data suggest, in fact, that voters are strategic in the way posited by such theory. For example, among voters who ranked Reagan first in 1980, 98% voted for their most preferred candidate, of those who most preferred Carter, 85% voted for Carter, but of those who most preferred Anderson, only 41% voted for him. Two attempts have been made to more carefully test for strategic behavior, using data from Britain (Cain 1978) and Canada (Black 1978). However, although expression (8) forms the theoretical basis for both of these studies, neither actually tests for the influence of the PB terms in ways prescribed by the model. Black uses the P terms alone whereas Cain drops some relevant terms (the utilities for the third candidate) without explanation. Thus, the interpretations of their empirical findings are blurred. Here, we intend to use the exact variables predicted to enter the utility calculation, namely, $P_{12}B_{12}$, $P_{13}B_{13}$ and $P_{23}B_{23}$.

Table 2 reports the estimation results for the Downsian model. The signs of the PB terms are exactly as predicted, with all significant above the 0.05 level. Thus, the expected utility terms that are unimportant in predicting turnout are highly significant when we change our dependent variable to the choice of a candidate. This finding should not, in retrospect, be surprising. When deciding whether or not to vote, the Downsian calculus assumes that a person compares the benefits of voting with the *cost* of voting. Given that the cost of voting is self-evident and given that a voter must pay this cost with certainty, any factor to be seriously considered as a “benefit” must be nearly certain as well. Hence, expected utility terms are unlikely to significantly affect a person’s decision.

¹¹Because third preference voting is *a priori* a dominated strategy, we only consider the vote choice between the first and the second preferences.

Rather, it is the social and psychological factors, the “consumption” value of voting, that, if anything, outweighs cost. On the other hand, when deciding for whom to vote, the voter’s calculation is based solely on the comparison of the net benefits of *voting for alternative candidates*, with both the D terms (including the P terms) and the C terms in the turnout calculation rendered irrelevant. In this event, only the expected utility terms are given a theoretical role, which is precisely what the data shows.

It appears, then, that a *rational voter* evaluates the decision to vote as a consumption act and the decision for whom to vote as an investment act. However, note that we have stacked the deck in favor of this interpretation of rationality. By including only PB terms in our estimated model we disallow the possibility that the candidate-choice decision is also a consumption act — model (8) is a pure “investment” model, in which variables like partisanship are not allowed to play an independent role. Research on American electoral behavior, though, demonstrates clearly that partisanship is a more stable factor than candidate evaluations and that it has its own independent influence on the vote decision (*cf.* Niemi and Weisberg 1984). Model (8), then, may misspecify the relationship between candidate evaluations and the decision for whom to vote so as to give us a false reading of the role of the PB terms.

The most straightforward approach to this problem is to add one or more terms to equation (3) or (7) in order to represent the *consumption value of voting for a specific candidate*. Denoting these terms by D_k , we can suppose, for example, that the more strongly the voter identifies with candidate k ’s party, the more utility the voter receives from voting for k . We can then rewrite equation (7) as:

$$E^k - E^0 = B_{k1}P_{k1} + B_{k2}P_{k2} + \dots + f(P_{ij}, \forall i \neq j) + D - C + D_k \quad (9)$$

and simple algebra yields the counterpart to model (8):

$$E^1 - E^2 = 2P_{12}B_{12} + P_{13}B_{13} - P_{23}B_{23} + D_1 - D_2 \quad (10)$$

For purposes of estimation, the new term $D_1 - D_2$ is operationalized by three new variables: the thermometer scores associated with the voter’s first and the second preference (U_1 and U_2 respectively), and the difference in the voter’s strength of partisan identification ($Party = Party_1 - Party_2$).¹² Naturally, we would expect that U_1 and $Party$

¹²The variables $Party_1$ and $Party_2$ are constructed from the 7 point party ID variable. Suppose a voter’s first preference is Reagan. Then $Party_1 = 4$ if the voter’s party ID is 6 (strong Republican); $Party_1 = 3$ if his party ID is 5 (weak Republican); $Party_1 = 2$ if party ID is 4 or 3 (independent leaning Republican and Independent) and $Party_1 = 1$ for all other party ID scores. The construction is similar if his first preference is Carter. If a voter’s first preference is Anderson, then $Party_1 = 4$ if party ID is 3 (Independent); 3 if party ID is 2 or 4 (Partisan independent); 2 if party ID is 1 or 5 (weak partisan) and 1 if party ID is 0 or 6 (strong partisan). $Party_2$ is obtained in a similar fashion. This construction measures both the direction and intensity of party ID relative to each preference.

positively affect first preference voting, whereas the effect of U_2 is negative.¹³ Our primary interest, though, is ascertaining what effect the inclusion of such variables has on our estimates of the coefficients for the PB terms.

Our results are reported in Table 3, and in comparison with Table 2 we see that the model significantly increases the log likelihood and that a likelihood ratio test shows that the modified model outperforms the original Downsian specification. More importantly, the significance of the PB terms are largely unaffected by the inclusion of U_1 , U_2 , and $Party$, all three of which have the right sign.

Are the right voters strategic? Uncovering evidence of strategic voting takes us only part way through the implications of model (10). What remains to be seen is whether the *right* voters are strategic. First, if a voter's first preference is a third-party candidate, then $P_{12}B_{12}$ and $P_{13}B_{13}$ should be near or at zero since the probability of first place tie between a third-party candidate and either of the two major contenders is, by definition, nil. Thus, voters who most prefer a third-party candidate should vote for their first preference only if:

$$E^1 - E^2 = -P_{23}B_{23} + D_1 - D_2 > 0$$

This expression formalizes the strategic voting hypothesis, which states that if the race between a second and last choice is close and if a voter likes the second choice much more than the last, then that voter should be more likely to vote for a second choice than for the first (Abramson, *et al*, 1992). Econometrically, this expression predicts that, among those voters who most prefer the weakest candidate, $P_{13}B_{13}$ and $P_{12}B_{12}$ should play an insignificant role, whereas the coefficient on $P_{23}B_{23}$ should be significant and negative. In contrast, for voters who most prefer a major party candidate, P_{23} should be close to zero and they should vote for their first choice if:

$$E^1 - E^2 = 2P_{12}B_{12} + P_{13}B_{13} + D_1 - D_2 > 0$$

Thus, a voter who most prefers a viable candidate — a candidate with a significant if not a certain chance of winning a plurality in that voter's state — should not act strategically and the estimated parameter associated with $P_{23}B_{23}$ should not be statistically significant.

Table 4 presents the relevant comparisons, where we assume that all voters favoring Anderson in 1980 are considered favoring a third party candidate, whereas for the 1968 (1992) election, those favoring Wallace (Perot) in states where Wallace (Perot) finished third are classified as such voters.

¹³Tests of multicollinearity show that the precision of U_1 and U_2 are affected, but all other parameters are reliable.

Limiting our attention to 1968 and 1980 for the moment, among voters favoring a major party candidate (“others”), the only variable significantly affecting the choice decision is party identification, with the PB terms playing no apparent role. Thus, as predicted, we can detect no strategic behavior among voters who most prefer a viable candidate. In contrast, PB_{23} clearly enters the voter’s calculus (with the proper sign) for voters favoring a third-party candidate. That is, as the race tightens between a voter’s second and third preference, as a voter’s preference for the second choice increases relative to that of the last choice, and to the extent that one of these candidates is the likely winner, then that voter becomes increasingly likely to cast a strategic vote for the second ranked candidate. We have here, then, clear evidence of strategic voting associated with the third party “squeeze” that is the basis of Duverger’s hypothesis about winner-take-all plurality elections.

Our results for 1992 are somewhat more ambiguous. Here PB_{13} is significant for those who rank the least competitive candidate (usually Perot) first, whereas PB_{23} is significant for “other” voters. We suspect, though, that we can attribute these results to misclassifications of “major” and “minor” party candidates: determining the “minor” party candidate on the basis of post-election returns within each state, we have misidentified the *subjective* evaluations of candidate viability for a significant number of respondents. Misidentification mixes “minor” and “major” party voters, thereby increasing the significance of variables that are predicted to be insignificant for one category of voters, and decreasing the significance of other variables predicted to be significant for a different category. This problem should arise with considerably less frequency for 1968 and 1980, owing to Anderson and Wallace’s clear standing as the least viable candidate. However, with Perot coming within 5 percent of Clinton in Kansas, 2 percent in Alaska, 1 percent in Idaho, 3 percent ahead of Clinton in Utah, and within 6 percent of Bush in Massachusetts, Nebraska, and Rhode Island, and ahead of Bush by three hundred and sixteen votes in Maine, there is much room, given the margin of error in polls generally, for voters to “misidentify” the viability rankings of candidates within a state. It is also the case that ranking of viability within a state need not correspond to their national ranking, and our analysis makes no effort at determining how voters might combine strategic calculations based on state-level measures of competitiveness with national considerations.

Nevertheless, despite these caveats, the fact remains that those expected utility terms that do not influence the decision to vote are significant when we explore the calculus of strategic voting.

4 Summary and Conclusion

After distinguishing between models in which closeness of the race alone enters a person's calculus of voting from the standard expected utility maximization model, we argue that the former are actually a form of pure "consumption" model in which closeness is *not* a probability discount factor but an independent variable bearing its own substantive meaning equivalent in effect to the "sense of citizen duty" postulated by Riker and Ordeshook (1968) to explain why people vote. This view not only provides a turnout model distinct from the Downsian one, but it also implies that empirical evidence establishing the relevance of the P term alone cannot be taken as support for the Downsian expected-utility model. After adding the minmax regret model postulated by Ferejohn and Fiorina (1974, 1975) to the analysis, we find that for the US presidential elections of 1968, 1980, and 1992, P but not PB or B can be verified as having a significant independent effect on the decision to vote.

It would appear, then, that support for the relevance of a rational choice analysis of voting is weakened by our analysis. Certainly, we cannot escape the conclusion that the decision to vote is largely a "consumptive" decision that is not based on a calculation about the likelihood of casting a decisive vote, and that the algebra commonly associated with rational choice models is largely irrelevant to understanding turnout. However, the turnout decision and the decision to cast a vote for a particular candidate are distinct. And the theoretical importance of expected utility terms differs from decision to decision, especially in three-candidate contests. Although probabilities (closeness) plays no theoretical role in two-candidate contests when we express a voter's candidate choice calculus, three candidate elections open the door to strategic calculations that depend on a voter's estimate of the viability of different candidates. This is the hypothesis that is the focus of our analysis and this is the hypothesis that is supported by the data. Specifically, in deciding for whom to vote, expected utility terms play a statistically significant role consonant with the concept of strategic voting.

How citizens act in an election is, of course, one of the most thoroughly studied topics in political science and is also one of the most controversial ones. "There are countless ways of understanding voting" (Niemi and Weisberg 1984). We have a full range of seemingly different and even opposing models: "social-historical" models, "sociological" models, "social-psychological" models, and "rational" models. However, the controversies are largely artificial. If utility is understood in a broad sense to include all types of fulfillment and satisfactions, then we can obtain a "unifying" model in which citizens can be described as maximizing this "gross" utility. At times this representation provides us with little that cannot be deduced or inferred from some other model, especially when we ask: What are the determinants of the different satisfactions that motivate people? Thus, for those circumstances in which decisions depend on strategically uncomplicated satisfactions – the decision to vote – a rational choice perspective provides little additional information about how people act. It is the "social-psychological" or "sociological" mod-

els – those that try to uncover the inputs into strategically uncomplicated satisfaction – that must take analytic precedence. On the other hand, when strategic considerations are thought to play a critical role, then the door must be opened to the possibility that the determinants of satisfaction influence decisions in more complicated ways – ways that can only be understood by theorizing carefully about what it means to be strategic.

Our findings confirm that, in the decision to vote, closeness of the race is a factor that independently contributes to the gross utility of voting but that expected utility calculations and the like add little to our understanding of events. It remains for proponents of “social- psychological” or other approaches to discover why this is so – why and how citizens are motivated to vote simply because an election is thought to be close. In contrast, the rational choice theorist has something specific to say when strategic calculation shapes the way different satisfactions influence decisions. Specifically, because the decision for whom to vote is strategically non-trivial in multi-candidate contests, simple models that merely “add up” satisfactions are likely to yield incorrect inferences. It is here that rational choice theorists can say something non-trivial about structural specification and thereby add to the research agendas of other approaches.

Appendix: Non-Nested Hypothesis Testing

Many of the familiar hypothesis testing situations involve *nested* models, *i.e.*, one model (the *restricted model*) can be obtained from the other model (the *unrestricted model*) under the null hypothesis about the parameters. The t test on the significance of a parameter of a model, for example, is a testing of nested hypothesis, with the unrestricted model being the model estimated, the restricted model being a special case of this model when one of the parameters is zero. Statistical testing techniques are well developed for this type of hypothesis testing. The classical Wald, Lagrange Multiplier and Likelihood Ratio tests are all for nested hypotheses. However, there are instances where we wish to compare two models, and one is not a special case of the other. The comparison of the various rational voter models is such a case. Neither can be obtained as a (parametric) special case of the other. Comparison of such models need to be carried out by *non-nested* hypothesis testing techniques. Non-nested hypothesis testing is not as mature as its nested hypothesis counterpart. However, we show below that one of such tests described by Ben-Akiva and Lerman (1985, p.171-172) can be simple to use and we produce a mini-table of the critical values of the test.

Suppose we need to compare model 1 and model 2 which are not nested. It was shown that, under the null hypothesis that model 1 is true, the following holds asymptotically:

$$Pr(\bar{\rho}_2^2 - \bar{\rho}_1^2 > z) \leq \Phi\{-[-2zL(0) + (K_2 - K_1)]^{1/2}\}, \quad z > 0 \quad (11)$$

where Φ is the standard normal cumulative distribution function, $\bar{\rho}_h^2$ is the adjusted likelihood ratio index for model $h = 1, 2$, $\bar{\rho}_h^2 = 1 - \frac{L_h(\hat{\beta}) - K_h}{L(0)}$, $L(0)$ is the initial log likelihood of the models, and $L_h(\hat{\beta})$ are the log likelihood at convergence of model $h = 1, 2$, K_h is the number of parameters of model $h = 1, 2$. Expression (12) means that the probability that model 1, which has the lower adjusted likelihood ratio index, is the true model is bounded by the right hand side of (12). Note that $\bar{\rho}_2^2 - \bar{\rho}_1^2 = \frac{[l(2) - l(1)] + [K_2 - K_1]}{l(0)}$, where $l(h) = |L_h(\hat{\beta})|$ and $l(0) = |L(0)|$. Hence the right hand side of (12) becomes: $\Phi\{-[2(l(1) - l(2)) + (K_1 - K_2)]^{1/2}\}$. Now suppose we want to find the critical values of the (two-tailed) test at significance level 0.05. Because $\Phi(-1.96) = .025$, we set $[2(l(1) - l(2)) + (K_1 - K_2)]^{1/2} = 1.96$, then:

$$l(1) - l(2) = 1.92 - 1/2(K_1 - K_2)$$

This means, for example, that if the two models have the same number of parameters, but model 2's log likelihood is more than 1.92 lower than model 1's in absolute values, then the probability that model 1 is the true model is less than 0.05. Table A lists the critical values for some commonly used significance levels and various $K_1 - K_2$ values.

Table 1-1
Comparison of Rational Turnout Models: the 1968 Data

Var.	MLE				<i>t</i> -Statistic			
	Model 1	Model2	Model3	Model4	Model 1	Model2	Model3	Model4
Const.	-3.18	-1.13	-3.16	-3.30	-6.23***	-1.04	-6.20***	-6.70***
PB12	2.19	—	—	—	0.89	—	—	—
PB13	0.65	—	—	—	0.31	—	—	—
PB23	0.96	—	—	—	0.29	—	—	—
P12	—	-7.03	—	—	—	-1.86*	—	—
P13	—	-7.80	—	—	—	-2.16**	—	—
P23	—	-8.23	—	—	—	-2.55**	—	—
B12	—	—	-0.50	—	—	—	-1.12	—
B13	—	—	0.78	—	—	—	2.26**	—
Alien.	1.79	2.20	1.56	2.05	3.09***	4.41***	2.61***	4.21***
Info.	0.50	0.50	0.50	0.50	6.96***	6.97***	6.93***	6.98***
Educ.	0.28	0.25	0.26	0.28	3.75***	3.43***	3.53***	3.86***
Party	0.31	0.31	0.31	0.31	4.01***	4.04***	4.05***	4.13***
LL at conv.:	-601.04	-596.98	-598.82	-601.45	Initial LL: -874.06		N = 1261	

Model 1: the Downsian model.
 Model 2: the P-model.
 Model 3: the minmax regret model.
 Model 4: D and C terms only.

* significant at .10 level.
 ** significant at .05 level.
 *** significant at .01 level.

Table 1-2
Comparison of Rational Turnout Models: the 1980 Data

Var.	MLE				<i>t</i> -Statistic			
	Model 1	Model2	Model3	Model4	Model 1	Model2	Model3	Model4
Const.	-1.75	-5.02	-1.92	-1.51	-2.98***	-3.04***	-3.19***	-2.80***
PB12	-3.31	—	—	—	-0.92	—	—	—
PB13	-1.51	—	—	—	-0.87	—	—	—
PB23	0.63	—	—	—	0.20	—	—	—
P12	—	12.03	—	—	—	2.10**	—	—
P13	—	12.90	—	—	—	2.27**	—	—
P23	—	13.12	—	—	—	2.35**	—	—
B12	—	—	-0.52	—	—	—	-0.93	—
B13	—	—	-0.31	—	—	—	-0.67	—
Alien.	1.23	0.93	1.60	0.78	1.58	1.47	1.94*	1.25
Info.	0.92	0.88	0.96	0.89	3.61***	3.49***	3.75***	3.55***
Educ.	0.26	0.27	0.27	0.27	2.64***	2.66***	2.73***	2.71***
Party	0.30	0.31	0.29	0.28	3.20***	3.26***	3.10***	3.08***
LL at conv.:	-417.87	-415.45	-417.43	-418.63	Initial LL: -467.87		N = 675	

Model 1: the Downsian model.
 Model 2: the P-model.
 Model 3: the minmax regret model.
 Model 4: D ans C terms only.

* significant at .10 level.
 ** significant at .05 level.
 *** significant at .01 level.

Table 1-3
Comparison of Rational Turnout Models: the 1992 Data

Var.	MLE				<i>t</i> -Statistic			
	Model 1	Model2	Model3	Model4	Model 1	Model2	Model3	Model4
Const.	-3.84	-10.27	-3.94	-4.10	-10.2***	-5.34***	-10.5***	-11.2***
PB12	.96	—	—	—	0.43	—	—	—
PB13	6.08	—	—	—	3.13***	—	—	—
PB23	2.51	—	—	—	0.84	—	—	—
P12	—	19.03	—	—	—	3.12***	—	—
P13	—	22.34	—	—	—	3.76***	—	—
P23	—	18.81	—	—	—	3.17***	—	—
B12	—	—	-0.23	—	—	—	-0.56	—
B13	—	—	0.90	—	—	—	2.57**	—
Alien.	1.26	1.93	1.38	2.01	2.83***	5.01***	3.04***	5.40
Info.	0.53	0.53	0.52	0.53	8.33***	8.42***	8.36***	8.48***
Educ.	0.38	0.36	0.38	0.37	8.78***	8.50***	8.85***	8.69***
Party	0.36	0.36	0.36	0.37	6.14***	6.18***	6.30***	6.44***
LL at conv.:	-971.72	-968.80	-974.29	-978.35	Initial LL: -1450.1		N = 2092	

Model 1: the Downsian model.
 Model 2: the P-model.
 Model 3: the minmax regret model.
 Model 4: D and C terms only.

* significant at .10 level.
 ** significant at .05 level.
 *** significant at .01 level.

Table 2
Strategic Voting: Estimation of the Downsian Model

Variables	MLE			<i>t</i> -Statistic		
	1968	1980	1992	1968	1980	1992
Const.	1.28	0.08	0.55	4.34***	0.23	2.85***
PB12	47.59	47.66	74.06	5.48***	2.89***	6.56***
PB13	11.28	56.98	21.53	2.35**	4.32***	4.72***
PB23	-12.57	-27.35	-22.35	-2.87***	-5.11***	-5.72***
Initial LL:	-583.63	-237.75	-885.15			
LL at Conv.:	-170.88	-97.55	-365.09			
N:	842	343	1277			

* significant at .10 level.

** significant at .05 level.

*** significant at .01 level.

Table 3

Strategic Voting: Estimation of the Modified Model

Variables	MLE			<i>t</i> -Statistic		
	1968	1980	1992	1968	1980	1992
Const.	-0.22	-1.65	0.18	-0.27	-1.36	0.34
PB12	27.49	24.55	69.73	1.93*	1.37	3.20***
PB13	5.13	54.03	21.52	0.90	3.75***	4.07***
PB23	-12.87	-29.95	-21.85	-2.68***	-4.78***	-5.13***
U1	5.78	4.47	-0.19	2.30**	1.98**	-.08
U2	-4.19	-2.10	0.74	-1.77*	-1.02	0.34
Party	0.37	0.33	0.34	4.19***	2.86***	6.05***
Initial LL:	-586.63	-237.75	-885.15			
LL at Conv.:	-159.38	-90.69	-345.12			
N:	842	343	1277			

* significant at .10 level.

** significant at .05 level.

*** significant at .01 level.

Table 4
Strategic Voting: Comparison of Voters
Favoring a Third Candidate and Other Voters

Var.	MLE						t-Statistic					
	1968		1980		1992		1968		1980		1992	
	W.-voters †	Others	A.-voters	Others	P.-voters	Others	W.-voters	Others	A.-voters	Others	P.-voters	Others
Const.	-1.17	0.83	-2.63	0.46	-1.14	1.71	-0.46	0.78	-1.82*	0.19	-1.45	2.12**
PB12	18.10	4.00	-74.89	-85.86	-10.77	69.13	0.13	0.10	-0.64	-0.97	-0.16	2.05**
PB13	22.60	-6.22	49.36	12.43	48.4	22.23	0.63	-0.44	1.30	0.43	2.74**	2.02**
PB23	-21.33	1.76	-23.78	23.49	-29.74	-38.77	-1.90*	0.10	-2.33**	0.28	-3.56***	-2.31**
U1	7.97	0.11	8.94	27.52	4.35	-0.88	1.12	1.42	2.14**	1.32	0.85	-0.17
U2	-5.07	-0.12	-5.90	-28.12	-2.29	-0.31	-0.71	-1.45	-1.46	-1.39	-0.44	-0.06
Party	0.43	0.46	0.21	0.57	0.27	0.44	1.55	3.85***	1.49	2.25**	3.52***	5.12***
Initial LL:	-44.4	-470.	-74.9	-163.	-219.	-666.						
LL at conv.:	-22.2	-99.9	-57.6	-28.9	-174.	-163.						
N:	64	678	108	235	316	961						
% correct class.:	85.9	95.1	73.2	95.8	70.9	94.5						

† voters who prefer Wallace most in states where Wallace finished third.

* significant at .10 level.

** significant at .05 level.

*** significant at .01 level.

Table A
Non-Nested Hypothesis Testing: Critical Values of ($l_1 - l_2$)

$\alpha \backslash k_1 - k_2$	-5	-4	-3	-2	-1	0	1	2	3	4	5
0.01	5.82	5.32	4.82	4.32	3.82	3.32	2.82	2.32	1.82	1.32	0.82
0.05	4.42	3.92	3.42	2.92	2.42	1.92	1.42	0.92	0.42	-0.08	-0.58
0.10	3.86	3.36	2.86	2.36	1.86	1.36	0.86	0.36	-0.14	-0.64	-1.14

α : significance level.

$k_{1,2}$: Number of Parameters in model 1, 2.

Formula: $l_1 - l_2 = c - (k_1 - k_2)/2$

where $c = 3.32$ for $\alpha = 0.01$; $c = 1.92$ for $\alpha = 0.05$ and $c = 1.36$ for $\alpha = 0.10$.

References

- Abramson, P.R., J.H. Aldrich, and D. Rhode. 1982. *Changes and Continuity in the 1980 Elections*. Congressional Quarterly Press.
- Abramson, P.R., J.H. Aldrich, P. Paolino, and D. Rhode. 1992. "‘Sophisticated’ Voting in the 1988 Presidential Primaries." *American Political Science Review*. 86:55-69.
- Ashenfelter, O. and Kelley, S. Jr. 1975. "Determinants of participation in presidential elections." *The Journal of Law and Economics*. 18:695-733.
- Ben-Akiva, M. and Lerman, S. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand*. The MIT Press.
- Black, J.H. 1978. "The multicandidate calculus of voting: application to Canadian federal elections." *American Journal of Political Science*, Vol.22, No.3, 609-638.
- Cain, B.E. 1978. "Strategic voting in Britain." *American Journal of Political Science*, Vol. 22, No.3, 639-655.
- Cohen, L.R. and Uhlaner, C.J. 1991. "Participation, closeness, and economic voting in the 1960 presidential election." Prepared for the 1991 Public Choice Society Annual Meeting, New Orleans, March 1991.
- Downs, A. 1957. *An Economic Theory of Democracy*. New York, Harper and Row.
- Duverger, M. 1967. *Political Parties*. London: Lowe and Brydone.
- Farquharson, R. 1969. *The Theory of Voting*. New Haven: Yale University Press.
- Ferejohn, J.A. and Fiorina, M.P. 1974. "The paradox of not voting: a decision theoretical analysis." *American Political Science Review*. 68:525-536
- Ferejohn, J.A. and Fiorina, M.P. 1975. "Closeness counts only in horseshoes and dancing." *American Political Science Review*, 69:920-925.
- Foster, C.B. 1984. "The performance of rational voter models in recent presidential elections." *American Political Science Review*. 78:678-690.
- Green, Donald P. and Ian Shapiro, 1994. *Pathologies of Rational Choice Theory*. New Haven: Yale University Press.
- Hansen, S., Palfrey, T.R. and Rosenthal, H. 1987. "The Downsian model of electoral participation: formal theory and empirical analysis of the constituency size effect." *Public Choice*. 52: 15-33.
- Johnston, J. 1984. *Econometric Methods*. New York: McGraw-Hill.

- Judge, G. et. al. 1985. *The Theory and Practice of Econometrics*. New York: Wiley.
- Katosh, J.P. and Traugott, M.W. 1982. "Costs and values in the calculus of voting." *American Journal of Political Science*, Vol.26, No.2:361-376.
- Kiewiet, D.R. 1979. "Approval voting: the case of the 1968 election." *Polity*, 12:170-181.
- Laver, M. 1987. "The logic of plurality voting in multi-party systems." In M.J. Holler(ed.): *The Logic of Multi-party Systems*. Copyright 1987 by Martinus Nijhoff Publishers, Dordrecht, The Netherlands.
- Ledyard, J.O. 1984. "The pure theory of two-candidate elections." *Public Choice*, 44:7-41
- Marx, B.D. and Smith, E.P. 1990. "Weighted multicollinearity in logistic regression: diagnostics and biased estimation techniques with an example from lake acidification." *Can. J. Fish. Aquat. Sci.*, Vol. 47, No.6, pp. 1128-1135.
- McKelvey, R.D. and Ordeshook, P.C. 1972. "A general theory of the calculus of voting." In J. F.Herdon and J. L. Bernd, eds., *Mathematical Applications in Political Science*, VI. Charlottesville: University Press of Virginia, pp. 32-78.
- Myerson, R.B. and Weber, R.J. 1988. "A theory of voting equilibria." *Discussion Paper NO. 782*. The Center for Mathematical Studies in Economics and Management Science Northwestern University, Evanston, Illinois 60208.
- Niemi, R.G. and Weisberg H.F. 1984. *Controversies in Voting Behavior*. Second edition. Congressional Quarterly Press.
- Ordeshook, Peter C. 1994. "The Spatial Analysis of Elections and Committees: Four Decades of Research," in J. Mueller, ed. *Handbook of Public Choice*, N.Y.: Basil-Blackwell.
- Palfrey, T.R. 1989. "A mathematical proof of Duverger's Law." In P.C. Ordeshook(ed.): *Models of Strategic Choice in Politics*. Ann Arbor, Michigan: University of Michigan.
- Palfrey, T.R. and Poole, K.T. 1987. "The relationship between information, ideology and voting behavior." *American Journal of Political Science*, Vol.31, No.3:511-530.
- Palfrey, T.R. and Rosenthal, H. 1985. "Voter participation and strategic uncertainty." *American Political Science Review*. 79:62-78.
- Riker, W. H. and Ordeshook, P.C. 1968. "A theory of the calculus of voting." *American Political Science Review*. 62:25-42.

- Rosenthal, H. and Sen, S. 1973. "Electoral participation in the French Fifth Republic." *American Political Science Review*. 67: 29-54.
- Schaefer, R.L. 1986. "Alternative estimators in logistic regression when the data are collinear." *J. Statist. Comput. Simul.*, 25:75-91.
- Schaefer, R.L., Roi, L.D. and Wolfe, R.A. 1984. "A ridge logistic estimator." *Communications in Statistics: Theory and Methods*. Vol. 13, No. 1. 99-113.
- Schwartz, Thomas. 1987. "Your Vote Counts on the Way it is Counted," *Public Choice*, vol 54: 101-21.
- Varian, H. 1984. *Microeconomic Analysis*. New York: W.W. Norton and Company.
- Zeng, Langche. 1993. *Individual Choice in Political Economy*, PhD Dissertation, Humanities and Social Science, California Institute of Technology.